Sonar: Collaborative Educational Playground

Roba Olana
Computer Science, NYUAD
roba@nyu.edu

Advised by: Yasir Zaki, Christopher Hoadley

ABSTRACT

There is a deficit of educational content that works across contexts. Students learn at different paces, have different starting points, different language capacities, learn from different curriculum etc. We believe the solution space lies in the nexus of collaboration across both consumption and production of resources. We built a platform that fosters a community of learners. The tool would provide a carousel of community-generated options for a given concept. We plan on expanding the use case of the product by leveraging dedicated strong communities and content to have an educational systems playground across nearly all current EdTech use cases.

KEYWORDS

educational technology, forum, user generated content, UGC, flutter, educational content

Reference Format:
Roba Olana. 2022. Sonar: Collaborative Educational Playground. In *NYUAD Capstone Seminar Reports, Fall 2022, Abu Dhabi, UAE.* 6 pages.

1 INTRODUCTION

People increasingly come together online to collaborate and produce massive niche content. We also know that people learn at different paces, have different starting points, different language capacities, different curriculums, etc. We believe the solution space lies in the nexus of collaboration across both consumption and production of resources.

Teachers can be thought of as customizing and localizing agents. In Lebanon, when schools were closed due to conflict, WhatsApp became an informal distance-learning app for students and teachers. ("WhatsApp School Is Now in Session" 2020) The app allowed educators to share their lessons using voice notes, videos, and photos in group chats. While Lebanon's school curricula are offered in French and English, many speak only Arabic fluently. WhatsApp's voice notes allow Lebanese teachers to switch between the three with ease. We saw similar stories come out of Zimbabwe, India, etc.

The lessons they are teaching, i.e. their customizations can be shared not just with their direct learners but broader learners of the curriculum. We can also have more people join the pool to customize. In the spirit of schoolhouse.world, students, or really anyone can teach and even make more relevant content.

A platform that aids in having variety/quantity at this scale, in itself, would be a big leap in improving how curiosity is fostered. But the benefits don't just come from having variety. The plurality of content created would mean coverage of varied curriculums, languages, and grade levels in addition to varied content types (videos, text, memes, interactive content) to meet learners as they are. Educational TikTok videos have shown us that there is education in the guise of distraction. (Balaji Srinivasan 2021) Imagine being able to swipe right until one finds that meme or game or explanation that clicks. We built a prototype of the the playground that enables that.

2 PRODUCT

2.1 Prototype

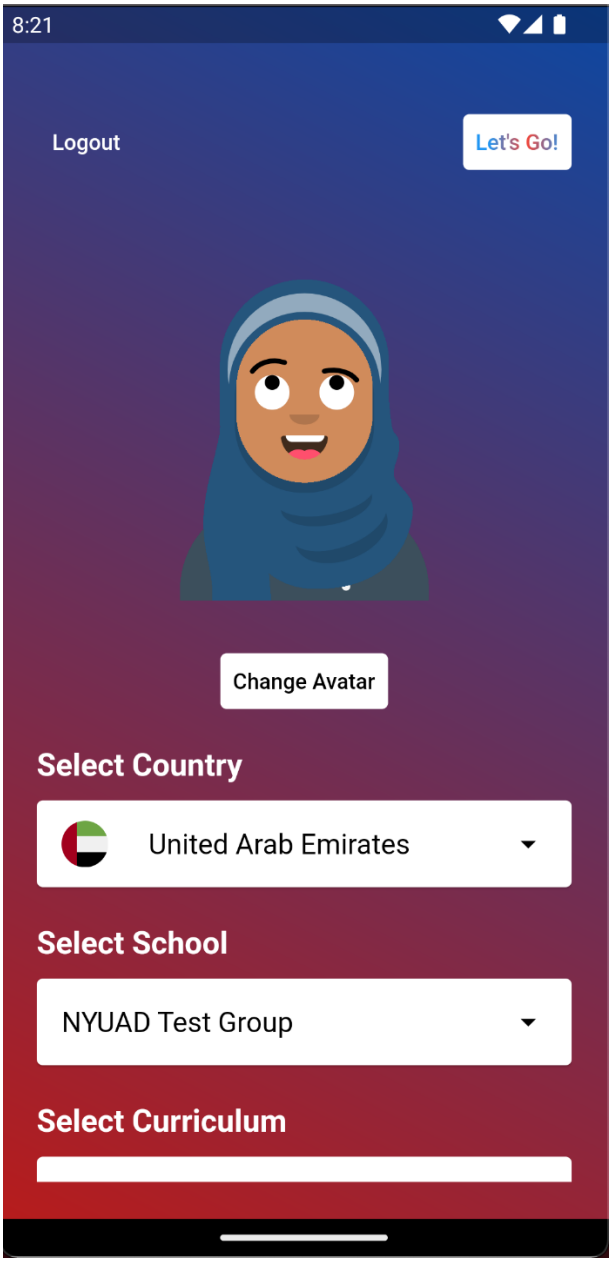
Increasingly getting to the big plan is key to building up a community to use this platform. The first phase of the project is focused on getting the community to produce and consume content based on an (initially) hardcoded curriculum graph. The curriculum is broken down to its most atomic units (concepts). For the user, this would mean a platform where they can select concepts and browse a variety of explanatory content. They can create Reddit like threads based on concepts and even specific content items.

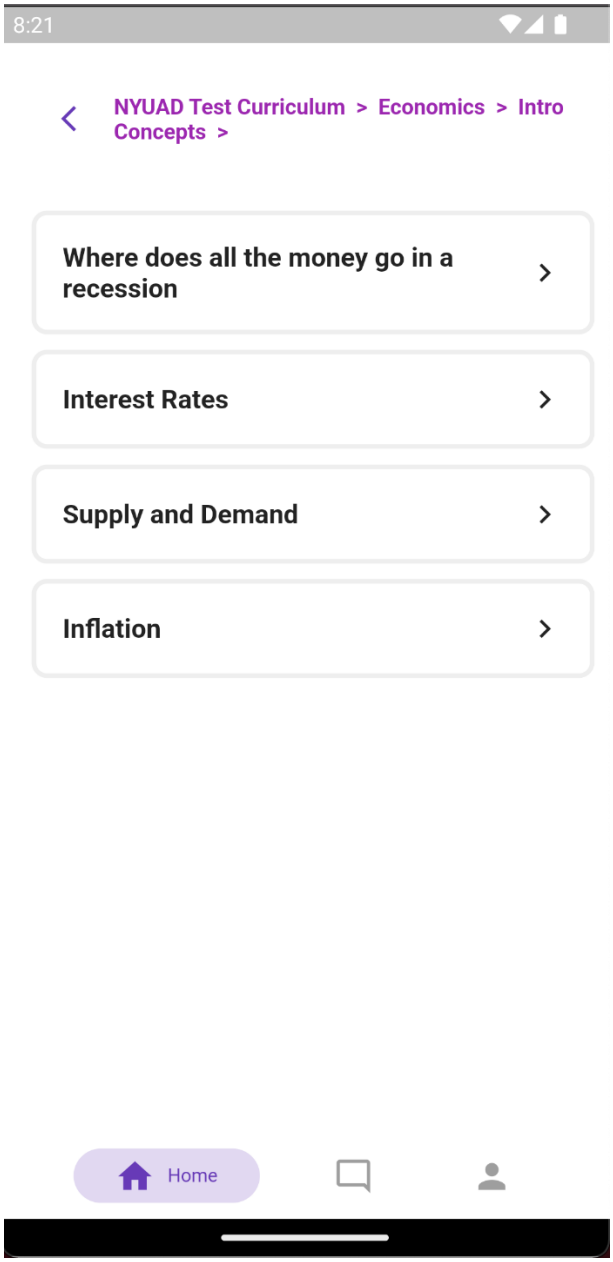
The forum is where the community gets to meet and rough edges get sanded out in the form of peer tutoring. User generated content is added in a carousel that can contain a video,

This report is submitted to NYUAD's capstone repository in fulfillment of NYUAD's Computer Science major graduation requirements.

جامعة نيويورك أبوظبي
NYU ABU DHABI

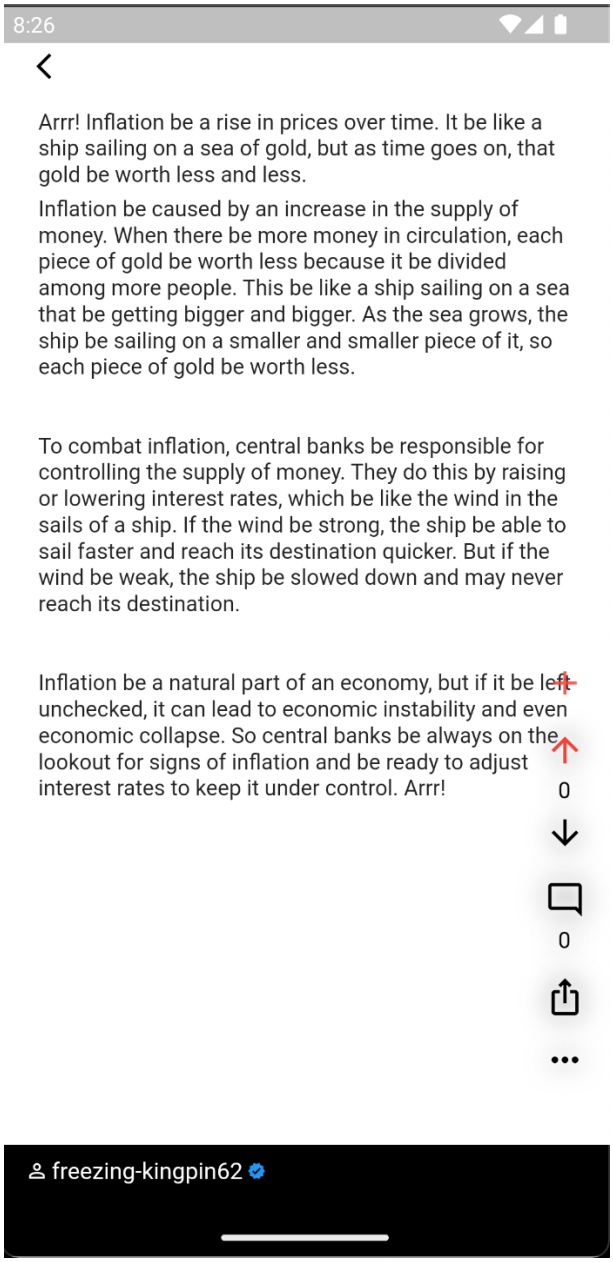
Capstone Seminar, Fall 2022, Abu Dhabi, UAE





image, audio, or a meme. The modular design of the application enables introduction of any content type using an SDK (Software development kit).

In producing the minimum viable product, the goal was to minimize development time and complexity to aid rapid experimentation. Application development SDKs are currently fragmented across operating systems. After analyzing leading cross platform application development frameworks, we opted for Flutter to develop the iOS, Android, and Web Application. Flutter provides its own UI components that build natively to each platform. It gives the benefit of native development while retaining high level development interoperability. In addition, it had one of the strongest development communities which meant a plurality of packages and online help. For the back end of the application, we opted for Firebase NoSQL storage by Google. We built data deserialization tools that convert data in the cloud to local classes for use by the application. To minimize costs, all database read and write are in the forum of futures (one time access, no update) with the exception of user reports which utilize streams (constant updates). User authentication is handled by Firebase Authentication APIs. While we have currently integrated



multiple sign in options, we only used email/phone numbers or school related login methods while interacting with beta testers.

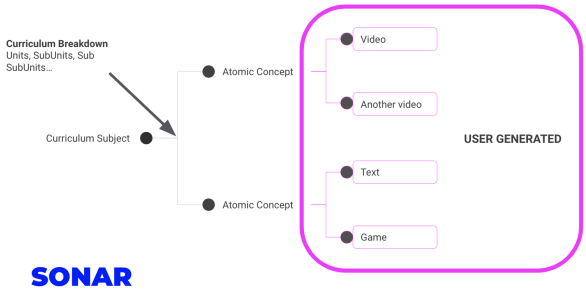
The forum is a semi-standalone part of the application that could be linked to concepts in the core application. It allows for multi nested commenting and upvote/downvote functionality. The concept items are separate widgets that display either images, video, audio or a combo of the reset in the form of an HTML file.

In the current MVP the users downvoting/upvoting content/comments/posts – serves to self-select for quality content. Having monetary incentives directly tied to content creation is a double edged sword (Gupta 2020) and we are thus initially going to seek only social returns for content creation with indirect monetary investments in community building. Our initial model is heavily influenced by stackOverflow moderation approach with the exception of letting users edit each other's posts. Barring success of this approach, we would be testing micropayment models with rates openly defined by the number of curriculum community members

and viewership. This we hope will create an ideal feedback loop where community members are both incentivized to create quality content and expand member count.

2.2 Assessment Layer

Given a community of learners and variety of content, we would be able to narrow down the state of the learner in relation to the knowledge space they are supposed to be in as defined by the curriculum. In order to do that, we would add an assessment layer and extensive semantic mapping out of curriculums. Imagine after scrolling right on options



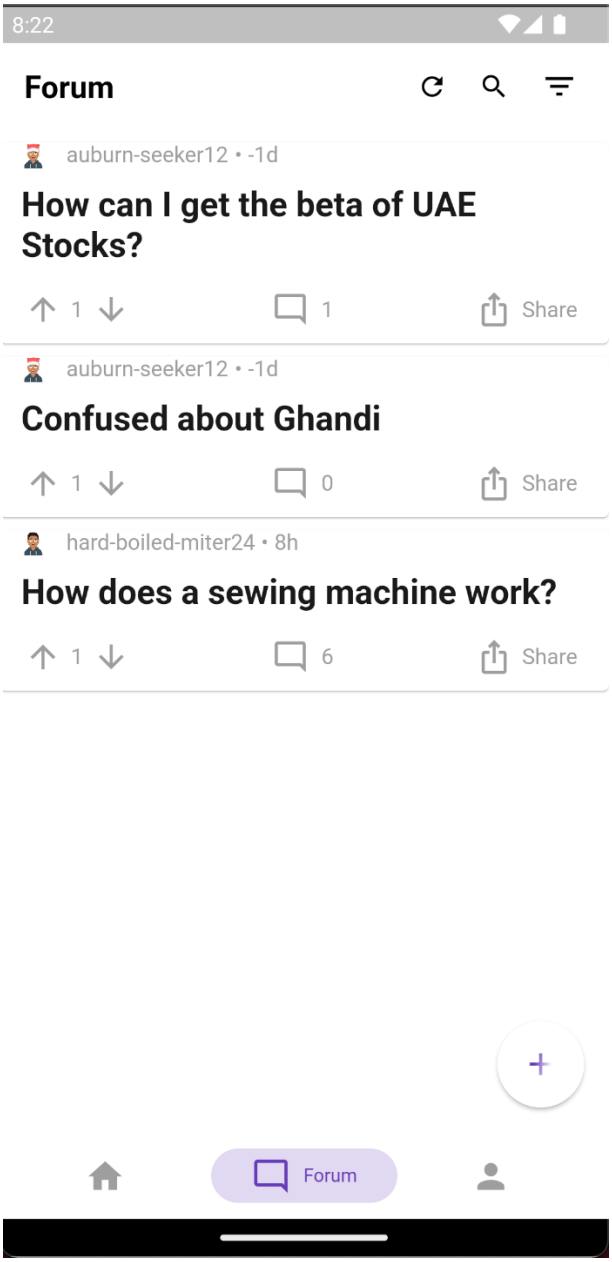
under a concept, scrolling up would reveal an assessment as part of an opt-in feature. Assessments are necessary in order to continually assess the learner on their knowledge space. (Falmagne et al. 2013) As such, instead of a linear scaffolding of content, an assessment layer that follows consumption of a concept examines mastery. Delivering on the promise of teaching at the right level, the system would be able to provide a skill from third grade to a ninth grade student to make sure they are up to par. For a given curriculum the complex relationship of skills and lessons can be represented in a knowledge graph. Let there be a student that is learning how to solve a word problem that requires a two-digit subtraction. A question might ask: "Jessica wants to buy a book that costs 15*andshehas*32 in her wallet. After she buys the book, how much money will she have left?"

Answering this question correctly requires a student to be able to perform a sequence of steps:

(1) Translate a word problem about subtraction into a mathematical expression: Question prompt $\rightarrow$ 32−15 = ?
(2) Solve the two-digit subtraction problem: 32−15
(3) Solve one-digit subtraction problems (while performing the two-digit sub-traction problem): 12−5, 2−1

These stages are concepts that can have their relationship defined. In the context of a knowledge graph (Wilson 2015), each step would be a skill represented as a node in the graph. A learner would have to know how to solve one-digit subtraction problems, solve two-digit subtraction problems, and translate a word problem into a mathematical expression. The figure below denotes the one directional relation between the skills graphically. Typically, a curriculum would be much more complex and more akin to the graph on the right.

After building these two layers, for the final stretch of tutoring we are looking at varied options ranging from simple Bayesian IRT algorithms or linking tutors directly, perhaps within the community. We also believe the output that comes out data collected as learners interact with content would be a massive driver for the learning engineering frontier. The end goal remains the same – the uniform syntax for learning content serves as a base for learning adaptation



systems to meet the needs of learners. It brings the benefits of adaptive tutoring to far more students and domains more quickly. While the rest of the platform is completely free, the assessment layer edges close to high stakes national exams which are closely tied to users' spending habits on EdTech. Besides tutoring, dedicated strong communities and content can be leveraged to have an educational systems playground across nearly all current EdTech use cases – that is where our moat is. For now, we are steadfastly focused on building the community using a minimum viable testing model.

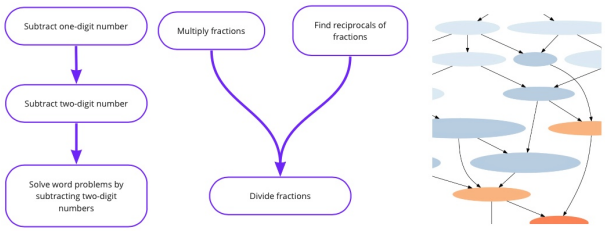


Figure 1: A simple mathematics knowledge graph (1), a knowledge graph with more than one prerequisite (2), a representation of a complex knowledge graph (3)

3 USER TESTING

There are three generally three ways an intervention such as this can be evaluated for its impact: use data about learning gains from test scores, make inferences based on usage metrics of the mobile application by learners, and use data about learning gains within the app to evaluate progression. For the scope of the current study, we opted to focus on evaluating the application development in the form of a usabilty testing.

Goal:

- Test if users can comfortably execute core application user scenarios
- Identify any critical and non-critical errors
- Gauge qualitative comfort of users in chosen application methodology

Quantitative Metrics:

- Successful Task Completion: Each scenario requires the participant to obtain specific data that would be used in a typical task. The scenario is successfully completed when the participant indicates they have found the answer or completed the task goal. In some cases, you may want give participants multiple-choice questions. Remember to include the questions and answers in the test plan and provide them to note-takers and observers.
- Critical Errors: Critical errors are deviations at completion from the targets of the scenario. For example, reporting the wrong data value due to the participant's workflow. Essentially the participant will not be able to finish the task. Participant may or may not be aware that the task goal is incorrect or incomplete.
- Non-Critical Errors: Non-critical errors are errors that are recovered by the participant and do not result in the participant's ability to successfully complete the task. These errors result in the task being completed less efficiently. For example, exploratory behaviors such as opening the wrong navigation menu item or using a control incorrectly are non-critical errors.

- Time On Task: The amount of time it takes the participant to complete the task.

Qualitative Metrics:

- Subjective Measures: These evaluations are self-reported participant ratings for satisfaction, ease of use, ease of finding information, etc where participants rate the measure on a 5 to 7-point Likert scale.
- Likes, Dislikes and Recommendations: Participants provide what they liked most about the site, what they liked least about the site, and recommendations for improving the site.

Approach:

- Sample size: Conventional wisdom from a commonly cited paper on usability test sample size postulates that to reveal 80 to 90 percent of interaction challenges, you should include five to eight participants per user group (Virzi et al, 1992).
- Participants were asked to complete task and their task completions and quantitative metrics were recorded

4 USER TESTING TAKEAWAYS

Critical Errors:

- Double clicking on a button while async processes didn't complete crushes the application
- Scrolling on a text page within a carousel stops user from going to the next concept

Non-critical Errors:

- Load times make content hard to view as the number of options increases, need to implement better pagination and preloading
- Videos don't pause when not in view sometimes
- Forum changes don't display unless application is refreshed

General Observations:

- Participants were more eager to product content for concepts subjects with higher potential viewership
- It's hard to consolidate a given 'concept' into just one video, image, text etc – platform should instead enable a 'storyline' type of presentation where for a given concept a given user can post blocks of content as a storyline that comprised of videos, (limited) text, code block, image etc.

5 CONCLUSION

We have found that there is a deficit of educational content that works across contexts. While individuals learn at different paces, have different starting points, different language capacities, learn from different curriculums etc, content available is often limited in its scope and contextual reach. We believe the solution space lies in the nexus of collaboration across both consumption and production of resources. We built a platform that fosters a community of learners. The user is able to swipe until they find a video, game or explanation that resonates. Initial user testing of the prototype are encouraging and indicate eagerness of learners to use a collaborative playground for learning.

REFERENCES

Barro, Robert J., and Jong Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950–2010." Journal of Development Economics 104 (September): 184–98. https://doi.org/10.1016/j.jdeveco.2012.10.001.

Balaji Srinivasan. 2021. "Education in the Guise of Distraction. Surprising Twists Courtesy of History." Tweet. @balajis (blog). July 8, 2021.

Evans, David K, and Amina Mendez Acosta. 2021. "Education in Africa: What Are We Learning?" Journal of African Economies 30 (1): 13–54. https://doi.org/10.1093/jae/ejaa009.

Falmagne, Jean-Claude, Dietrich Albert, Christopher Doble, David Eppstein, and Xiangen Hu, eds. 2013. Knowledge Spaces. Berlin, Heidelberg: Springer Berlin Heidelberg.

Gupta, Debjit. 2020. "Essays on the Management of Online Platforms: Bayesian Perspectives," August.

https://vtechworks.lib.vt.edu/handle/10919/99603.

"Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty: Banerjee, Abhijit V., Duflo, Esther: 9781610390934: Amazon.Com: Books." n.d. Accessed December 18, 2021.

serena li here. 2021. "@gaganbiyani @udemy Has a Very Special Place in My Heart. Institutional Edu Perhaps Shouldn't Be Free, but Also Shouldn't Be THAT Expensive" Tweet. @SerenaLiHere (blog). March 3, 2021.

UIS, UNESCO. 2016. "The World Needs Almost 69 Million New Teachers to Reach the 2030 Education Goals." November 17, 2016. http://uis.unesco.org/en/document/world-needs-almost-69-million-new-teachers-reach-2030-education-goals.

"WhatsApp School Is Now in Session." 2020. Rest of World. August 27, 2020. https://restofworld.org/2020/lebanon-schools-whatsapp-group/.

Wilson, Kevin. 2015. "A General-Purpose Adaptive Learning Infrastructure," 24.

"World Development Report 2018: Learning to Realize Education's Promise." n.d. World Bank. Accessed December 18, 2021. https://www.worldbank.org/en/publication/wdr2018.

Virzi, Robert A.. "Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough?" Human Factors: The Journal of Human Factors and Ergonomics Society 34 (1992): 457 - 468.